The Role of Empathy in the Recognition of Vocal Emotions

Rene Altrov, Hille Pajupuu, Jaan Pajupuu

Institute of the Estonian Language, Tallinn, Estonia

`rene.altrov@eki.ee, hille.pajupuu@eki.ee, jaan@pajupuu.eu`

Abstract

The study addresses the role of empathy in the recognition of vocal emotions (joy, anger, sadness) when they are moderately expressed in speech which is listened to without the speaker being seen. The test was taken by 67 adults (29 male and 38 female subjects aged 30–60), whose empathy level had previously been measured by Baron-Cohen & Wheelwright's self-report questionnaire, the Empathy Quotient (EQ). In the test group only men (n=14) had a low empathy level (EQ<32), and only women (n=16) had a high empathy level (EQ>51). A comparison of the men (n=15) and women (n=22) who had a medium empathy quotient showed that gender was a factor in the recognition of emotions. To eliminate the possible gender influence, the recognition of emotions was tested by comparing the results of the men with a low EQ with those of the men with a medium EQ and the results of the women with a medium EQ with those of the women with a high EQ. According to the results, personal empathic ability does not affect the recognition of emotions from the voice.

Index Terms: emotional speech, emotion recognition, empathy, empathy quotient, gender

1. Introduction

Empathy is considered important in social communication as it enables people to understand the emotions and affective states of others (cognitive empathy), to feel/experience them and to respond to them isomorphously or appropriately (affective empathy) [1]-[4].

Empathy has hitherto not received much attention from speech technologists, but over recent decades the focus of speech research has moved to emotions [5], which has raised a need for specific information on whether empathy may affect the recognition of emotions from the voice only, without the speaker being seen. Speech technology studies often use testers or evaluators whose task is to recognize, classify or mark emotions in human or synthetic speech, but it is not known if the test results also depend on their empathic ability. The present study aims to find out if emotion research requires the empathy quotient of the testers/evaluators to be measured and taken into consideration in the interpretation of the results of emotion recognition.

The available studies on the relations of empathy and the ability of recognizing vocal emotions purely from the voice serve a mainly clinical purpose, by investigating how emotions are recognized by people whose diagnosis involves a low empathy level, e.g. autism, Asperger's Syndrome (AS), psychopathy. To our knowledge the influence of empathy on the recognition of vocal emotions has not been studied on non-clinical population, and thus our assumptions about such influence can be based on clinical results only. Notably, children with AS have been observed to have problems in the recognition of the basic emotions of joy, anger, and sadness from acted speech: in comparison with typically developing peers without a psychological diagnosis they determined more

prosody scenes as being neutral [6]. Asperger adolescents, however, did not differ from their typically developed peers in tests where happiness, sadness, anger and neutrality had to be recognized from texts read aloud by actors with no disturbing factors present [7]. Both groups did very well, recognizing all basic emotions with an accuracy of approximately 82%, while emotions were mostly confused with neutrality by both groups.

When it comes to complex emotions, however, the Asperger or autistic adults have encountered problems [8]-[10]. A low empathy level is negatively correlated with the ability to comprehend the affective state of another person. All the tests were based on acted emotional speech (segments of speech taken from BBC drama series and from audio books). A meta-analysis of the available studies have also revealed deficiencies in the recognition of emotions by children, adolescents and adults with psychopathic traits [11]. Specifically, psychopathy is associated with significantly poorer recognition of fear, happiness and surprise, but not of anger, disgust, or sadness.

Thus, most of the clinically oriented works have shown that low empathy is often accompanied by difficulties in the recognition of emotions and mental states from the voice.

Our study differs from the above-referenced ones in two aspects: (1) the subjects have no diagnosed disorders such that might have a reducing effect on empathy; (2) the emotions that need to be recognized are elicited, not acted. In addition, those are not full-blown basic emotions expressed stereotypically, which seldom occur in everyday communication, but moderately expressed emotions belonging to three groups: joy, anger, and sadness.

Our research question was whether a person's empathy level (low, medium, high) can affect their ability to recognize emotions from the voice.

2. Method

2.1. Material

The material comes from the Estonian Emotional Speech Corpus[1]. The examples of emotional speech in the corpus were obtained from longer text passages read aloud. Assuming that every text will evoke a certain mood that is expressed in the reader's voice, the read passages were segmented into sentences, which were separated from their context and presented to 30 listeners (evaluators). The evaluators were asked to identify the emotional tone of each sentence and respectively classify the sentences into four groups: joy, anger, sadness, or neutral. The instructions pointed out that each group also covered related similar emotions. In addition, the same sentences were presented to the evaluators in the written form and they were asked to determine the emotion from the linguistic content, so that they could try and guess the

---

[1] http://peeter.eki.ee:5000/

emotions without hearing the voice. The aim was to ascertain those sentences in which the emotion was detectable purely from vocal cues, without any contribution from the linguistic content [12].

The material of the present study consisted of 35 sentences (emotion identification rate no less than 65%), picked from the corpus, where the emotional attitude had been found to be accessible without knowing the linguistic content. In terms of emotions the sentences could be grouped under the categories of joy (10), sadness (10), anger (10), and neutral (5). The sentences were arranged into a web-based listening test. In sequencing the material it was kept in view that no two consecutive sentences would make a logical whole.

2.2. Participants and procedure

The listening test was taken by 67 adult volunteers (29 male and 38 female subjects aged 30–60, M=42.0, SD=10.7). Their empathy level was measured by means of an Estonian translation of a self-report questionnaire Empathy Quotient (EQ) designed by Baron-Cohen and Wheelwright [2]. This questionnaire is considered a reliable and valid tool to measure empathy in both healthy individuals and clinical populations [1], [13] and [14]. The EQ was accessed by 60 items, 40 of which addressed empathy, while 20 were filler items, which were included to distract the participant from a relentless focus on empathy. Each item was provided with four response options: "definitely agree", "slightly agree", "slightly disagree", and "strongly disagree". Approximately half the items were reversed, anticipating disagreement-responses from empathic subjects and agreement-responses from non-empathic ones. The subjects would score zero (0) points for a "non- empathic" response, regardless of its strength, and 1 or 2 points for an "empathic" response, depending on the strength of the reply. The maximum score was 80 points. According to the EQ test results, men tend to score less, on average, than women, while women with a low empathy level are few, and so are men with a high empathy level [2], [13]. The original test items and procedure are described and available in [2].

Firstly, the participants of the listening test were divided into three groups by their empathy scores: low empathy – under 32 points, medium empathy – 32–51 points, and high empathy – over 51 points, cf. [15].

Of our sample, only men belonged to the low-empathy group and only women had high empathy. Subsequently those with a medium empathy score were divided into male and female subgroups.

Test groups:

I. Low empathy: *EQ* 18–31, *M*=26.2, *SD*=3.2

    14 men aged 32–44, *M*=36.7, *SD*=4.5

II. Medium empathy: *EQ* 32–51, *M*=42.5, *SD*=5.7

    a) 15 men aged 31–56, *M*=39.0, *SD*=7.6; *EQ* 32–50, *M*=41.5, *SD*=5.8

    b) 22 women aged 31–60, *M*=38.5, *SD*=6.4; *EQ* 34–51, *M*=43.1, *SD*=5.1

III. High empathy: *EQ* 52–60, *M*=54.9, *SD*=2.6

    16 women aged 30–58, *M*=44.1, *SD*=10.5

The participants were then asked to listen to separate context-free sentences without seeing the text, and to decide what emotion was involved in each of them. The choice was among joy, anger, sadness, and neutrality. The instructions pointed out that as full-blown emotions are seldom met in normal speech, "joy" should be interpreted as also covering gratitude, happiness, pleasantness and exhilaration; while "anger" included resentment, irony, reluctance, contempt, malice and rage; and "sadness" covered loneliness, disconsolation, concern and hopelessness. "Neutral" meant normal speech without special emotions. The sentences could be listened to as many times as the participants wished.

2.3. Analysis

As the test groups differed in gender – including an exclusively male group (of low empathy) and an exclusively female group (of high empathy) – we first had to check the possible influence of gender on the recognition of vocal emotions. For that purpose we analysed the men and women with medium empathy as the only group containing both sexes. Pearson's chi-square test of independence was applied to evaluate the difference between the groups' abilities in emotion recognition (the difference is significant if p<0.05). The same method was used for pairwise comparison of groups with different empathy levels (men with low vs. medium empathy, women with medium vs. high empathy). A confusion pattern is presented to demonstrate the accuracy of the recognition of the target emotion in groups and with what emotions the target emotion was confused.

3. Results

Mutual comparison of men and women with medium empathy demonstrated gender influence on the recognition of two emotions: joy and anger (p<0.05) (see Table 1). Therefore the emotion scores were only compared within gender, comparing the male groups (low vs. medium empathy) with each other and so were the female groups (medium vs. high empathy). The results showed that there was no significant difference between the emotion recognition scores of the men from the low-empathy group and the men from the medium-empathy group, nor was there any significant difference in emotion recognition between the medium-empathy and high-empathy women (see Tables 2 and 3).

All the test groups recognized the target emotion of the test sentences with an accuracy of over 50% (in most cases more than twice as high as chance probability), see Table 4.

Table 1. *Gender influence on the recognition of vocal emotions: Comparison of men and women having a medium empathy level*

Target emotion	Groups EQ 32–51	Responses				*p*
		Joy	Anger	Sadness	Neutral	
Joy	men	96	9	12	41	.001
	women	182	9	13	27	
Anger	men	6	83	26	42	.001
	women	5	174	22	33	
Sadness	men	1	24	115	15	.123
	women	2	26	194	11	
Neutral	men	6	13	10	50	.113
	women	8	7	16	87	

Note. Pearson's χ2 results. Difference in the recognition of an emotion is significant if *p* < .05.

Table 2. *The influence of empathy on the recognition of vocal emotions: Comparison of male groups with low and medium empathy*

Target emotion	Groups EQ 18-44	Responses				p
		Joy	Anger	Sadness	Neutral	
Joy	low	116	9	12	39	.809
	medium	96	9	12	41	
Anger	low	7	114	33	34	.275
	medium	6	83	26	42	
Sadness	low	1	21	146	10	.346
	medium	1	24	115	15	
Neutral	low	13	12	18	49	.267
	medium	6	13	10	50	

Note. Pearson's χ2 results. Difference in the recognition of an emotion is significant if *p*<0.05.

Table 3. *The influence of empathy on the recognition of vocal emotions: Comparison of female groups with medium and high empathy*

Target emotion	Groups EQ 32-60	Responses				p
		Joy	Anger	Sadness	Neutral	
Joy	medium	182	9	13	27	.096
	high	132	10	10	38	
Anger	medium	5	174	22	33	.241
	high	8	124	21	35	
Sadness	medium	2	26	194	11	.338
	high	5	15	158	11	
Neutral	medium	8	7	16	87	.165
	high	9	9	5	73	

Note. Pearson's χ2 results. Difference in the recognition of an emotion is significant if *p*<0.05.

Table 4. *Confusion matrix: Recognition of Estonian emotions by men and women with a low, medium or high EQ (% of the recognition of the target emotion)*

EQ	Target emotion	Response emotions			
		Joy	Anger	Sadness	Neutral
52-60 high	**Women**				
	Joy	**69.5**	5.3	5.3	20.0
	Anger	4.3	**65.9**	11.2	18.6
	Sadness	2.6	7.9	**83.6**	5.8
	Neutral	9.4	9.4	5.2	**76.0**
32-51 medium	**Women**				
	Joy	**78.8**	3.9	5.6	11.7
	Anger	2.1	**74.4**	9.4	14.1
	Sadness	0.9	11.2	**83.2**	4.7
	Neutral	6.8	5.9	13.6	**73.7**
	Men				
	Joy	**60.8**	5.7	7.6	25.9
	Anger	3.8	**52.9**	16.6	26.7
	Sadness	0.6	15.5	**74.2**	9.7
	Neutral	7.6	16.4	12.7	**63.3**
18-31 low	**Men**				
	Joy	**65.9**	5.1	6.8	22.2
	Anger	3.7	**60.6**	17.6	18.1
	Sadness	0.6	11.8	**82.0**	5.6
	Neutral	14.1	13.0	19.6	**53.3**

4. Discussion

The aim of the study was to find out how one's empathic ability might possibly affect the recognition of emotions if the decision has to be made from nothing but the voice. The results showed that people with high, medium and low empathy levels did not significantly differ in their scores of emotion recognition (Tables 2-3). In the test sentences, where emotions were moderately expressed, joy, anger, sadness, and a neutral attitude were recognized in more than 50% of cases, irrespective of the subject's gender or empathy level (Table 4).

These group results seem to differ from those of the previous studies, which have demonstrated the difficulty of non-empathic people in the recognition of vocal emotions, see [6], [8]-[11]. However, several conditions of our study also differed from those of the previous ones. First, our study was not clinically oriented: the low-empathy subjects had not been diagnosed to have empathy-reducing conditions or developmental disorders. They were normal adults typically developing without psychiatric disorder. Secondly, the test material was different. In the previous studies the participants were to listen either to acted basic emotions or to sentences/paragraphs conveying complex emotions, which had been extracted from TV drama series or audio books. Our study did not use actors. Instead, the listening material consisted of sentences read by ordinary people, while their moderately expressed emotions were elicited by the emotional content of a longer text passage. No full-blown emotions were manifested in the test material. Although the perceived emotion was to be classified under joy, anger, sadness or a neutral attitude, the classes were also meant to cover comparable emotions; e.g. joy was the generic emotion for gratitude, happiness, pleasantness etc. Thirdly, the testing conditions were different: our participants had access to the listening test on the web and they could listen to the sentences as many times as they wished, without any time restrictions.

Considering the above differences in the material, participants, and testing conditions, our results will neither confirm nor disprove the previous results. Instead, they supplement additional knowledge to the effect that ordinary empathic and non-empathic adults do not significantly differ in the recognition of moderately expressed elicited vocal emotions if they have sufficient time.

This result can be useful in teaching emotions to autistic people who find emotions difficult to interpret. Considering the progress achieved in emotion recognition by training autistic people on acted emotions without any time limit for interpretation[2] [9], the next step "closer to life" should train them to understand moderately expressed non-acted emotions. Our results based on healthy people suggest that training autistic people could be efficient if they have enough time for interpreting the emotions or if they can listen to the vocal expressions repeatedly.

Another result of the present study was related to the difference between male and female subjects in emotion recognition: men and women with a medium empathic ability differed in the recognition of emotions belonging to the classes of joy and anger (see Table 1), and therefore our investigation of the influence of the empathic ability on emotion recognition had to be conducted on men and women separately. Table 4 reveals, that the male subjects whose empathy was from low to medium got lower scores for the recognition of the target

---

[2] MIND READING: The Interactive Guide to Emotions http://www.jkp.com/mindreading/faq.php/

emotion than the female subjects, whose empathy was from medium to high.

Generalizing the obtained results, it can be concluded that in the recognition of vocal emotions, empathy is less relevant than suggested by clinical studies. Consequently, when studying emotions for speech technological purposes it is hardly necessary to exclude non-empathic people from testers or evaluators for fear that they might fail to recognize emotions, if only their low empathy level is not due to psychiatric or developmental disorders. It would be expedient, however, to use both male and female subjects as they differ in their ability in emotion recognition.

There are also certain limitations to our study. First, according to the results of the empathy test, the low-empathy group consisted of men only and the high-empathy group consisted of women only. Thus we could neither compare how emotions are recognized by low-empathy men vs. low-empathy women, nor the relevant scores of high-empathy men vs. high-empathy women. As, in a general case, women tend to be more empathic than men anyway [2], [13], it is hard to recruit comparable groups. Second, our test enabled repeated listening to the test sentences. On the one hand, it has given us the knowledge that non-empathic people can recognize emotions in sentences heard several times, but on the other hand, we do not know whether the results would have been the same if the emotion had been decided upon in real time.

We agree with Dziobek [4] that in future studies of the influence of empathy on emotion, recognition should be continued in situations which are as authentic as possible, where the voice is supported by the visual aspect. Also, more research attention could be given to the influence of personality traits on emotion recognition, which may be found to be even more important than empathy.

5. Conclusions

Empathy has been a topic of interest for different fields of research, such as neuro-psychology, developmental psychology, philosophy, anthropology etc. Our interest was inspired by speech technology, raising the question: what is the role of empathy in emotion recognition? Our results show that in a test of emotion recognition, where the emotion is conveyed by the voice only (no mimics or bodily movement can be seen) and no time restriction is applied, the empathic ability does not significantly affect emotion recognition. Hence we conclude that speech technology studies requiring understanding, analysis and measurement of emotions need not particularly worry about the empathic ability of the participants.

6. Acknowledgements

7. References

[1] C. Allison, S. Baron-Cohen, S. Wheelwright, M. H. Stone and S. J. Muncer, "Psychometric analysis of the Empathy Quotient (EQ)", *Personality and Individual Differences*, vol. 51, no. 7, pp. 829–835, 2011.

[2] S. Baron-Cohen and S. Wheelwright, "The Empathy Quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences", *Journal of Autism and Developmental Disorders*, vol. 34, no. 2, pp. 163–175, 2004.

[3] H. Walter, "Social cognitive neuroscience of empathy: Concepts, circuits, and genes", *Emotion Review*, vol. 4, no. 1, pp. 9–17, 2012.

[4] I. Dziobek, "Comment: Towards a more ecologically valid assessment of empathy", *Emotion Review*, vol. 4, no. 1, pp. 18–19, 2012.

[5] J. Tao and T. Tan, Eds., *Affective Information Processing*. London: Springer, 2009.

[6] J. L. Lindner and L. A. Rosén, „Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger's syndrome", *Journal of Autism and Developmental Disorder*s, vol. 36, no. 6, pp. 769–777, 2006.

[7] J. Heikkinen, E. Jansson-Verkasalo, J. Toivanen, K. Suominen, E. Väyrynen, I. Moilanen and T. Seppänen, "Perception of basic emotions from speech prosody in adolescents with Asperger's syndrome", *Logopedics, Phoniatrics, Vocology*, vol. 35, pp. 113–120, 2009.

[8] M. D. Rutherford, S. Baron-Cohen and S. Wheelwright, "Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning autism", *Journal of Autism and Developmental Disorders*, vol. 32, no. 3, pp. 189–94, 2002.

[9] O. Golan, S. Baron-Cohen and J. Hill, "The Cambridge Mindreading (CAM) Face-Voice Battery: Testing complex emotion recognition in adults with and without Asperger syndrome", *Journal of Autism and Developmental Disorders*, vol. 36, no. 2, pp. 169–83, 2006.

[10] O. Golan, S. Baron-Cohen, J. Hill and M. D. Rutherford, "The "Reading the Mind in the Voice" test-revised: A study of complex emotion recognition in adults with and without autism spectrum conditions", *Journal of Autism and Developmental Disorders*, vol. 37, no. 6, pp. 1096–1106, 2007.

[11] A. Dawel, R. O'Kearney, E. McKone and R. Palermo, "Not just fear and sadness: Meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy", *Neuroscience and biobehavioral reviews*, vol. 36, no. 10, pp. 2288–2304, 2012.

[12] R. Altrov and H. Pajupuu, "Estonian Emotional Speech Corpus: Theoretical base and implementation", in L. Devillers, B. Schuller, A. Batliner, P. Rosso, E. Douglas-Cowie, R. Cowie, C. Pelachaud [Eds.], *Proc. of the 4th Int. Workshop on Corpora for Research on Emotion Sentiment & Social Signals (ES3)*, pp. 50–53, Istanbul, 2012.

[13] E. J. Lawrence, P. Shaw, D. Baker, S. Baron-Cohen and A.S. David, "Measuring empathy: Reliability and validity of the empathy quotient", *Psychological Medicine*, vol. 34, pp. 911–924, 2004.

[14] S. J. Muncer and J. Ling, "Psychometric analysis of the empathy quotient (EQ) scale", *Personality and Individual Differences*, vol. 40, pp. 1111–1119, 2006.

[15] S. Bate, B. Parris, C. Haslam and J. Kay, "Socio-emotional functioning and face recognition ability in the normal population", *Personality and Individual Differences*, vol. 48, no. 2, pp. 239–242, 2010.